Detecting Faces from Color Video by Using Paired Wavelet Features

Szu-Hao Huang and Shang-Hong Lai
Computer Vision Laboratory,
Dept. of Computer Science,
National Tsing Hua University, Hsinchu, Taiwan 300
{ Howard, lai }@cs.nthu.edu.tw

Abstract

Detecting human face regions in color video is normally required for further processing in many practical applications. In this paper, we propose a learning-based algorithm that determines the most discriminative pairs of Haar wavelet coefficients of color images for face detection. To select the most discriminative features from the vast amount (1,492,128) of possible pairs of three-channel color wavelet coefficients, we employ two procedures to accomplish this task. At first, we choose a subset of effective candidate pairs of wavelet coefficients based on the Kullback Leibler (KL) distance between the conditional joint distributions of the face and non-face training data. Then, the adaboost algorithm is employed to incrementally select a set of complementary pairs of wavelet coefficients and determine the best combination of weak classifiers that are based on the joint conditional probabilities of these selected coefficient pairs for face detection.

By applying Kalman filter to predict and update the face region in a video, we extending the face detection from a single image to a video sequence. In contrast to the previous face detection works, the proposed algorithm is based on finding the discriminative features of joint wavelet coefficients computed from all three channels of color images in an integrated learning framework. We experimentally show that the proposed algorithm can achieve high accuracy and fast speed for detecting faces from color video.

1. Introduction

Most automatic face-related applications and researches require a robust face detector as the first step. These applications, including face recognition, face tracking, expression cloning, face post estimation, and 3D head model reconstruction from images, usually assume the face region is localized very correctly. Sometimes, they require the detection of facial features as the next step.

In recent years, a number of machine learning techniques have been applied to the face detection problem. Most previous face detection methods are based on using gray-scale features only for classification. Ignoring color information may lose important information, thus making the face detection problem very difficult. Some of them use skin color filters as a pre-filter to skip some impossible regions to speedup the face detection. Unfortunately, the wide varieties of lighting conditions make the independent skin color filters or grayscale face detectors very difficult. This cascaded approach implicitly assumes the independence of the color and brightness. It does not fully exploit the brightness and color information for face detection. In this paper, we focus on finding the most discriminative color features for face detection by using statistical learning techniques. Then, we combine this color face detector with Kalman filter to detect faces from video.

To determine the most discriminative wavelet features for face detection, we search all pairs of the three-level wavelet coefficients from the three-channel $Y\text{-}C_b\text{-}C_r$ color face and non-face images. The 2D face and non-face distributions of each wavelet coefficient pair are estimated from all training examples to determine the discrimination power of each pair of wavelet coefficients. For a 24-by-24 training image, which only contains 576 pixels, the total number of possible wavelet coefficient pairs from the three-level wavelet transforms of three channels grows to 1,492,128. It is very difficult for traditional training algorithms to handle such a high-dimensional feature space for a limited training time and memory. By combining some statistical learning techniques, we develop an efficient face detector based on combining discriminative pairs of color wavelet coefficients. To find the powerful pairs of wavelet coefficients for face detection, we compute the Kullback Leibler (KL) distance between face and non-face distributions for each pair of wavelet coefficients to select a subset of candidate wavelet

Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)
1063-6919/04 $ 20.00 IEEE

coefficient pairs. This feature screening technique is used to effectively reduce the total number of candidate wavelet coefficient pairs for face detection. Subsequently, we compute the discrete joint probability density function of all the candidate pairs of wavelet coefficients and apply the adaboost training algorithm to select a set of complementary pairs of wavelet coefficients as well as the weights of the linear combination of the corresponding weak classifiers to be the final face classifier. Then, we combine the above face classifier with Kalman filter to predict and refine the estimate of face location in video. The Kalman filter not only provides the estimation of the mean and covariance of the face location parameters but also gives prediction for the face location statistics in the next frame, which can be used to determine a suitable search area of faces. This idea of reduced area search by using Kalman filter prediction leads to a very fast real-time face detection system for a video.

2. Previous Works

Most previous face detection methods focused on detecting faces from a single gray-scale image. The survey paper [6] by Yang et al. classified the face detection methods into four categories; namely, knowledge-based methods, feature –based methods, template matching methods [8], and appearance-based method.

The appearance-based approach has evolved to be a major stream in the face detection research. Since it is very hard to describe a general face in an image by some explicit characterization or feature description, the appearance-based approach learns to determine a face through face and non-face examples. The training stage in this approach is to decide a two-class classifier from training examples. After colleting a large number of training face images, some researchers focus on finding a suitable classifier for face detection.

Statistical modeling techniques have been applied to model face images in a reduced- dimensional feature space. The eigenface approach [7] employed the principle component analysis to find the most representative eigenvectors and reduce the feature dimension. The example-based approach [1] estimated the distribution of face and non-face training samples in a high-dimensional feature space by using mixtures of Gaussians. The neural network approach [3] adopted the normalized intensities in overlapping regions as the input features. Later, they generalized this method [4] to allow planar rotation of face images.

Support Vector Machine (SVM) has been successfully applied to face detection [9]. The SVM classifier seeks the kernel hyperplane with the largest margin between positive and negative training data. Heisele et al. [5] combine the wavelet coefficients with SVM to develop a robust face detection system. Recently, they have advanced their method [10] to component-based analysis to detect faces of different poses.

Recently, Viola and Jones [2] proposed a real-time face detection algorithm, which combines many weak classifiers based on selecting simple features that can be computed from the integral image very easily. After adaboost training and bootstraping, the face classifier can work in real time with high accuracy. Our proposed algorithm is also based on combining weak classifiers via adaboost training, but there are fundamental differences between these two algorithms, especially in the use of color wavelet coefficient pair features and their selection.

Most previous methods used the color information to design skin-color filters as the first step of face detection [18], followed by a gray-scale face detector. Huang and Lai [19] also employed a skin-color filter as a screening pre-filter and used a Gaussian location model to achieve real-time face detection from video. In addition, Hsu et al. [17] employed color features to detect eye and mouth regions. A distinguishable feature of the proposed algorithm is the selection of discriminative pairs of color wavelet coefficients as the features for face detection through a statistical learning process.

3. Proposed Face Detection Learning Method

Our proposed face detection system is based on learning from examples. We collect a large number of low-resolution color face images as training data and apply some machine learning techniques to find the powerful wavelet features for our face detector.

The features used in our face detection system are discriminative pairs of color wavelet coefficients. The training algorithm focuses on effectively finding discriminative features from the training data. In our system, a novel quantization procedure and a learning-based feature selection scheme are used to find the most powerful features for face detection. With the proposed learning algorithm, the total number of features used for face detection is reduced, thus reducing the time required for the adaboost training and face detection.

The final face detector is a linear combination of the outputs from a set of weak classifiers, which determine if a region in an image is a face from the selected pairs of color wavelet coefficients. Note that the final selection of complementary pairs of wavelet coefficients for the weak classifiers and the weights of their linear combination are learned through the adaboost training.

3.1 Haar Wavelet Coefficient Pair Features

A distinguished characteristic of the proposed face detection algorithm is the rich information contained in a very limited set of color wavelet feature. This attribute is derived from the paired combination of multi-resolution wavelet coefficients in three color channels.

In our feature computation, we first convert the color images to the Y-C_b-C_r color space, which is commonly used in color images compression. For each image, we compute three-level wavelet transformations for the three color channels, respectively. Since Haar wavelet transform can be computed with additions and subtractions only, this makes the wavelet feature be computed very rapidly in the detection process. Figure 1 shows the representation of wavelet coefficients.

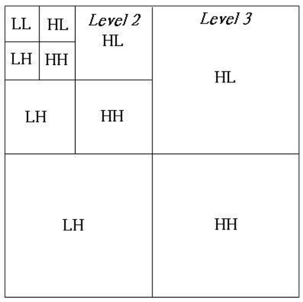


Figure 1. Wavelet representation of an image**.**

For a 24-by-24 three-channel color image, we have totally 1,728 wavelet coefficients. Then the total number of all possible pairs of these wavelet coefficients is 1,492,128. Different types of wavelet coefficient pairs are discussed below. Each pair of wavelet coefficients belongs to three categories from the following six categories.

1. Inter-color channels – Coefficient pairs from different color channels may represent information between the space and color. This type of features includes the joint event of red lip and white teeth.

2. Intra-color channel – These features focus on the face structure information. It is similar to the traditional gray-level information, but we have more varieties in the pairs corresponding to the C_b and C_r channels. In order to prove our training is effective, we have experimented only in Y inter-channel features and compare the results with some previous systems. The result is shown in section 5.1.

3. Inter-wavelet levels – Across different levels, the features can represent the multi-resolution property. For example, we can describe the whole face and mouth symmetry jointly in one feature. This can complement the inadequate spatial relationship in the traditional features.

4. Intra-wavelet level – In the same wavelet level, the feature pairs represent the position relationship. Aided by color and different bands, these features still have much power to discriminate between faces and non-faces. For example, the relationship between left and right eyes can be found in this kind of feature.

5. Inter-wavelet bands – Combinations of wavelet coefficients from different bands bring us different meaning. For example, in the same color and level, the HL and LH pairs can represent the corner information. These features have more varieties for arbitrary combinations.

6. Intra-wavelet band – The same wavelet band has the same attributes. It can represent the spatial, scale, and color relationship in the same kind of features. For example, the eye and mouth regions have stronger horizontal edges than vertical edges.

3.2 Learning System Overview

The large number of using all possible pairs of wavelet coefficients as the features for face detection makes the training infeasible. To take advantage of the promising property of the wide varieties of feature representation, we must find a way to effectively reduce the large number of features and find a small set of discriminative features, i.e. the wavelet coefficient pairs, to construct our face detector.

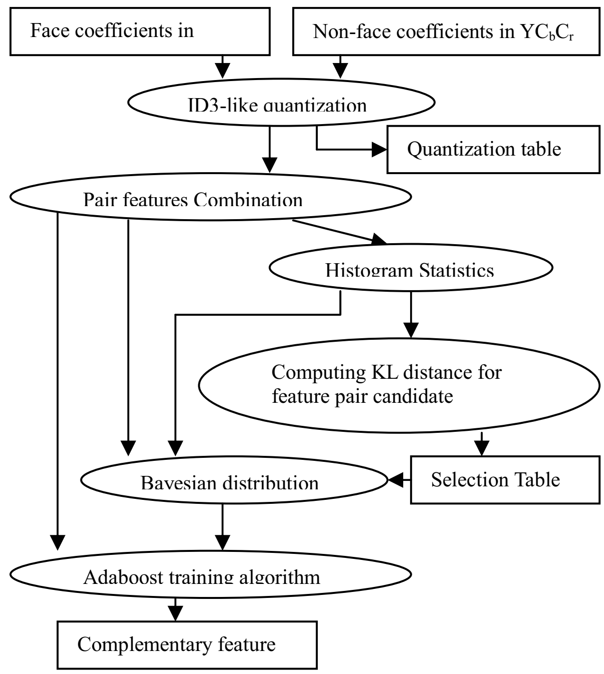


Figure 2. The flow chart of proposed training system

Figure 2 depicts the flow chart of the training system for our face detection algorithm. The proposed training system can be roughly divided into two stages. The first stage selects a subset of discriminative pairs of wavelet coefficients from all possible pairs by comparing the corresponding KL distance, which measures the distinction between the face and non-face distributions. This is followed by the adaboost algorithm, which selects a complementary set of features from this subset and determines a linear combination of weak classifiers as the final face detector. The details of each component in the training system are described in the subsequent sections.

3.3 ID3-like Balance Tree Quantization

This training system starts with an ID3-like balance tree quantization for computing the discrete joint

Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)
1063-6919/04 $ 20.00 IEEE

distributions corresponding to the wavelet coefficient pairs. This quantization assigns a wavelet coefficient into one of eight levels based on its distribution from the training data, with the aim of preserving the most discriminative information. This method works better than histogram equalization or fixed interval quantization.

In order to speed-up the quantization process, we select all the boundary seeds first. The best quantization boundaries certainly appear between two ordered values corresponding to positive data and negative data. This idea of boundary seed selection is sketched in Figure 3.

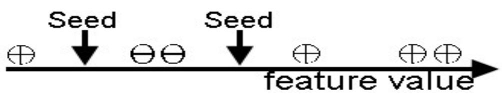


Figure 3. Seeds Selection Sketch

The main algorithm of ID-3 decision tree is to select the best boundary in each node which can divide the data passing this node into two classes with largest information gain. It means that the selected boundary can help each branch contain data of the same class as much as possible. In other words, we want to find appropriate boundaries to distribute data into intervals of maximal uniformity.

In the ID3-decision tree, we first define the entropy and information gain as follows:

$$\text{Entropy}(S) = -p_{\oplus}\log_2 p_{\oplus} - p_{\Theta}\log_2 p_{\Theta} \qquad (1)$$

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in leaf\_nodes} \frac{|S_v|}{|S|} Entropy(S_v) \qquad (2)$$

where $p_{\oplus}$ and p_{Θ} are the probabilities of positive and negative samples in the data set S, respectively, the symbol A denotes a threshold to divide a set S into two subsets S_v. Then, we can select the best seed value that maximizes the information gain as follows:

$$selected\_seeds = \arg\max_{A}(Gain(S, A)) \qquad (3)$$

Figure 4 shows the structure of a balanced decision tree quantization.

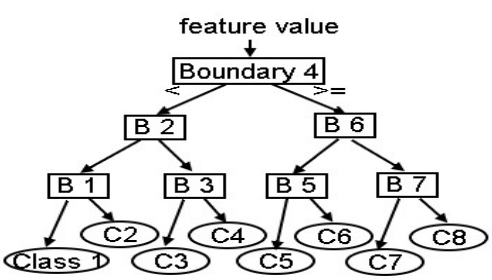


Figure 4. Balanced Tree Quantization.

3.4 Kullback-Leibler Discrimination Measure

The Kullback-Leibler (KL) distance is used to measure the discrimination between the distributions of positive and negative training data. In each mapped feature plane, the wavelet coefficient pairs with larger KL distances indicate that they are more suitable to distinguish the patterns between these two classes. It is reasonable to assume that most of the wavelet pairs lack of discrimination power, therefore it is imperative to select the discriminative pairs as the candidate features for face detection.

The KL distance between the face and non-face distributions is used to select candidates of paired wavelet features, thus making our learning algorithm more efficient. The KL distance is computed from the distributions of positive and negative data in the 2D feature pair space. It is given as follows:

$$D_{KL} = D(P(pos) \,\|\, P(neg)) + D(P(neg) \,\|\, P(pos)) \qquad (4)$$

where

$$D(P(pos) \,\|\, P(neg)) = \sum_{i \in all\_intervals} P_i(pos) \ln \frac{P_i(neg)}{P_i(pos)} \qquad (5)$$

3.5 Bayesian Weak Classifiers

For each pair of wavelet coefficients, we can train a weak classifier based on this paired feature. The AdaBoost training algorithm is then used to select some powerful weak classifiers and combine them to determine if it is a face. For each weak classifier, we apply the Bayesian decision rule to measure the face conditional probability density function in each interval. By thresholding the ratio of face and non-face probability for a given paired feature vector, we have a weak classifier that can be used in the AdaBoost training algorithm.

Making a hard decision in the weak classifier may lose information that was computed in the conditional probabilities. We want to use the conditional probability directly into the AdaBoost training algorithm. The experimental results show this probabilistic weak classifier works better than binary weak classifier and the convergence speed increases in the AdaBoost training.

By applying the Bayes rule, we can compute the conditional probability as follows

$$\frac{p(pos \mid X)}{p(neg \mid X) + p(pos \mid X)} = \frac{p(X \mid pos)}{p(X \mid pos) + p(X \mid neg) \cdot \dfrac{p(neg)}{p(pos)}} \qquad (6)$$

Equation (6) outputs a value between 0 and 1, which represents the face conditional probability.

3.6 Adaboost Training Algorithm

In comparison with other kinds of training algorithms, adaboost has a unique attribute that it selects discriminative features recursively during the training process. For speed consideration, it is important to select a small number of discriminative features and determine whether it is a face very quickly. Although checking all features can bring higher accuracy, we must compromise between accuracy and execution speed.

The adaboost learning algorithm is used to achieve the selection of complementary weak classifiers and determine the associated weights at the same time. Combination of the best few classifiers may not work better than combination of the same number of weak classifiers with complementary attributes [2, 14]. The details of the Adaboost algorithm are referred to [14]. In this paper, we modified the algorithm by replacing the binary weak classifier to the conditional probability output.

Iterative bootstrap training has been employed in many face detection systems. We also adopt this training scheme in our system training. After the adaBoost training from a set of training images to produce the first set of weak classifiers, we apply the combined face classifier to detect faces in a large set of natural images that contain no faces. The detected face windows must be false alarms and are fed into the training kernel as new non-face training images. This process is repeated until convergence.

The final classifier trained by the adaboost algorithm is a weighted linear combination of the selected weak classifiers as given in equation (7).

$$FD(x) = \begin{cases} \text{true} & \text{when} \sum_{t=1}^{T} \alpha_t p(f_t(x)) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ \text{false} & \text{otherwise} \end{cases} \tag{7}$$

where x is a input test image, T is the total number of combinational features, α_t is the weight associated with the t-th weak classifier, $f_t(x)$ is the interval where x is located on the corresponding wavelet feature pair plane, the function p is the joint probability. Note that α_t and $f_t()$ are trained from the adaboost algorithm. The joint probability p is obtained from the associated Bayesian weak classifier.

3.7 Training Implementation

We collected 8600 low-resolution, 24-by-24 color face images as the training data. Some of them are depicted in Figure 5. The images are collect from two main sources: one is manually cut from Internet images containing face regions (2600 faces), and the other is cut from 1200 portraits from different people with slight rotation after manually alignment (6000 faces). The first part of the face training images contains more representation of different lighting situation and skin color variety. The second part focuses on diversified pose variations by rotating the face images with ± 5 and ± 10 degrees.

Figure 5. Training face image samples

The negative training data is more than one thousand natural images with sizes ranging from 320-by-240 to 1280-by-1024, mainly from Corel database. In the first step, we randomly select 4600 negative windows as the first set of non-face training data. Then, the new negative training set is produced by selecting the regions that pass the existing face detector. The reason we only pick 4600 non-face images in each iteration of training is based on the consideration of limited system memory and computation time.

In our implementation, we have three sets of training face images and three detectors of different scales. They are used to overcome the face scaling between the fixed scales provided by the multi-resolutional wavelet domain. This avoids the image downsampling process and repeated wavelet transformation, thus saving the computation power and execution time.

4. Face Detection in Color Video

In this section, we describe the details of applying the color face detector to detect faces from a video. In our system, we first detect faces from a single image with our adaboost color face detector, then the Kalman filter is applied to update the detected face location and predict the new location in the next frame. In the next frame, the color face detector is applied to the areas determined from the Kalman filter prediction.

Before the first face appears in the video frame, we check all possible regions in a video frame with our color face detector. After simple lighting normalization and Haar wavelet transformation, we compute the face conditional probabilities of all the selected weak classifiers through table look-up and compute the weighted combination to be the final strong face detector result. Because of our effective training system, we can achieve high accuracy with only a small set of discriminative feature pairs. The performance between accuracy and false positive will be illustrated in the experimental section.

In addition, slight modification of face detection structure also reduces much time for detecting faces of different scales. We trained three face detectors, as described in section 5.7, which covered different face scales between consecutive levels in wavelet transform. We only need to compute the wavelet trasnform once per input frame, thus it is needless to downsample the image and compute the wavelet transformation again.

After first detecting faces from video frames, we try to model the position and motion of face region in the video frames. Kalman filter can iteratively update the mean and covariance matrix of face location based on previous face detection estimates and predict the position in the next frame. In our Kalman filter implementation, the state vector consists of six elements which are the position in x and y coordinates, face region size, velocity in x and y coordinates, and the temporal change speed of face scale.

We use the position predicted by Kalman filter as the detection center. Assuming the x, y, scale to be independent, we use the diagonal elements of covariance matrix to determine the search area range. For example, if the variance of face scale is larger than others, we will find more different scales in the neighbor regions.

5. Experimental Results

The experiments to be shown in this section can be divided into three parts. The first one is a simulation that compares the result between the gray-scale version of our method on gray-scale images and some previous distinguished face detectors. The second part is still a simulation with a fixed training and testing database. We will show how much our system progresses when color information is used. The third part is face detection from real world video sequence.

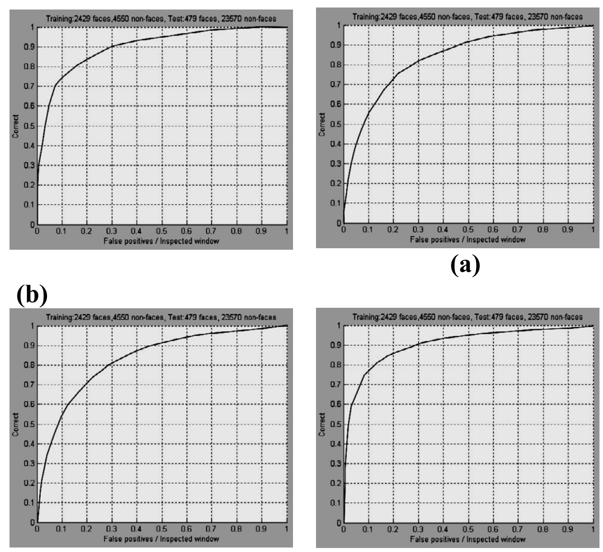
5.1 Comparison in Gray-scale Image

This experiment is to prove our training system is effective. Although our system is based on color images, it still can be modified to train and test gray-scale images. The modified version loses the color variety attribute, but it still keeps the restriction of wavelet domain. We show the result is as well as, or even better than, some of the best gray-scale face detection systems.

MIT CBCL face data set provides us a fair environment to compare the classifier performance. The training data set contains 6977 images (2429 face and 4548 non-face) and testing data set is composed of 24045 images (472 face and 23573).

Our experimental result as shown in the ROC curves in Figure 6 can be reproduced by any other person who wants to examine our performance. And the result of SNoW-based method and SVM method are referred to [5][15][16].

Just like linear kernel SVM, our system is based on linear combination of features. Although we eliminate line fitting preprocessing and replace histogram equalization by simple normalization, we can still obtain better result than the linear kernel SVM approach. Moreover, our system can classify all training images correctly. The other two systems are much more different designs. It is clear that we perform better than SNoW-based face detection system on this data set. Compared with the polynomial-kernal SVM method, our method can not attain as high detection rate as theirs at lower false positive, but our curve outperforms their result when false positive is larger than 0.5. The difference between our proposed method and polynomial-kernal SVM are very small. The characteristics of these two methods may bring different benefits to different kinds of applications.



(a)

(b)

(c)

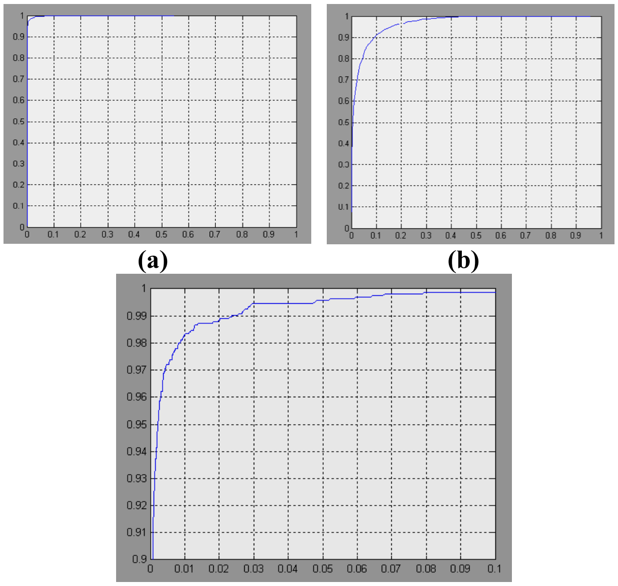
(d)

Figure 6. ROC curves of different systems (a) our proposed method (b) linear kernel SVM (c) SNoW-based system (d) polynomial-kernel SVM

5.2 Improvement by Color Information

The CBCL is just a gray-scale image data set. We can not judge the performance of our final detector with color information on the database. We additionally clip 328 faces of different people that are not included in the training database. After slightly rotated ±5, ±10 degree, the face testing database has 1640 face images. Besides, randomly extracting 23,724 non-face regions from Corel database brings us a wide variety of negative testing set.

We want to prove the progress of color information and judge our system performance when the feature variety increases. The detectors in this experiment are produced after first iteration. Without iteratively training, we already can get quite nice simulation result. Our experiments on color and gray level images here use the same training and testing images. The gray-scale detector only acquires Y color channel information and extracts intra-color features. The color detector is trained as we discussed before.



(a) **(b)** **(c)**

Figure 7. The ROC curves of (a) our color system, (b) our gray-scale system, and (c) the enlarged ROC curve of our color system

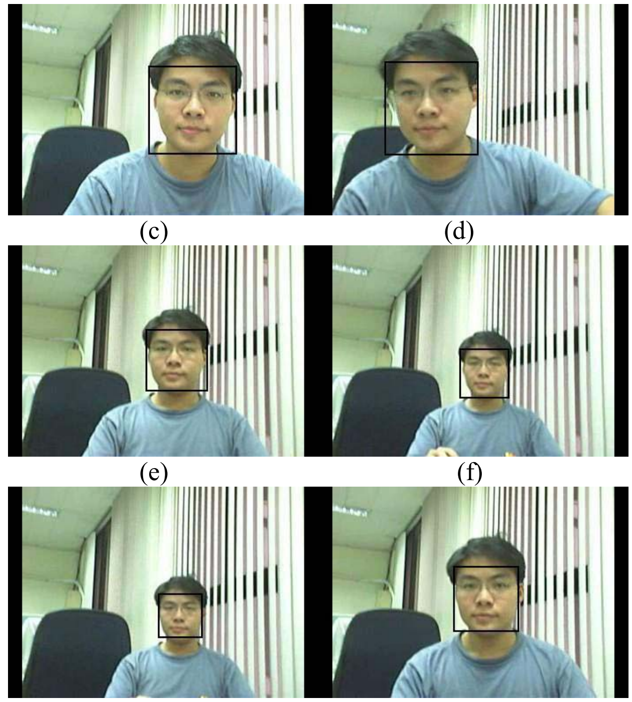
It is obvious the color information improves the face detection performance dramatically. Large feature variety brings more rich information for us to train the most discriminative classifier.

5.3 Detecting face in video

For the face detection from video, we depict some detection results on two sets of video frames. The video is captured from web camera in the lab.

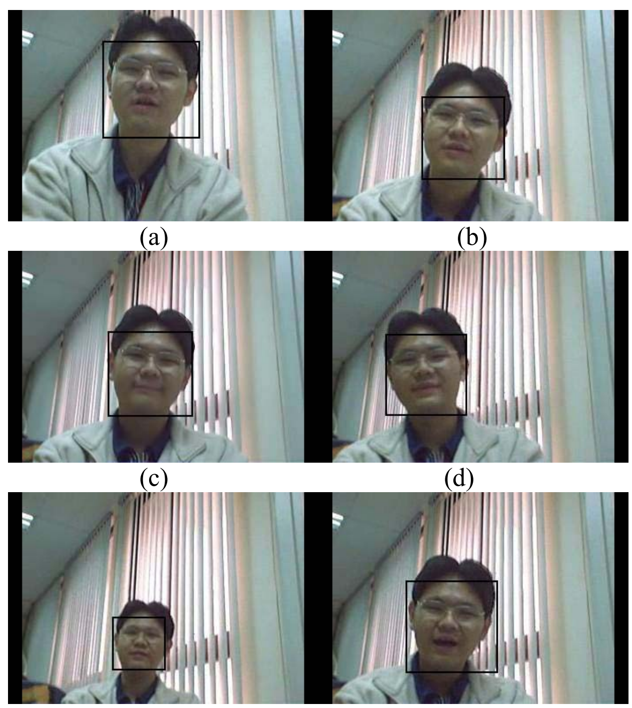
The frames we displayed here are randomly selected from the testing sequence. Our face detector can indicate face regions correctly from all frames of our testing video as shown in Figure 8 and 9.

(a) (b)



(c) (d) (e) (f) (g) (h)

Figure 8. The frames extracted from detection result of video Howard.mpg where (a) is frame no. 12 (b) 110 (c)253 (d) 300 (e) 335 (f) 396 (g) 442 (h) 595



(a) (b) (c) (d) (e) (f)

(g) (h)

Figure 9. The frames extracted from detection result of video Chung.mpg where (a) is frame no. 1 (b) 51 (c)101 (d) 151 (e) 201 (f) 251 (g) 301 (h) 351

6. Discussion and Conclusion

This paper presented a novel face detection algorithm from color images in a wavelet domain. Aided by the rich information provided by the wavelet feature pair combination, we can make use of the relationship between color, position, and scale learned from training color face images. The main idea of our design followed two concepts. The first one is preserving most information and finding the most discriminative pairs of wavelet coefficients to form the strong classifiers. The second idea is delicate training architecture and efficient execution on the wavelet domain. The KL distance is used to reduce the large number of wavelet pairs to a small subset. The Bayesian classifier provides the paired feature a trustworthy probability density function as a weak classifier. Then the AdaBoost algorithm selects a set of complementary weak classifiers. Bootstrap training is used to refine our classifier to reduce the false alarms from various kinds of non-face samples.

Kalman filter is used to extend the face detection algorithm from a single image to a video sequence. By taking advantage of the temporal model and prediction capability, we can reduce the search region and save computational time. In the future, we aim to improve this system to detect faces with large pose variations.

Acknowledgements

This research was jointly supported by the Program for Promoting Academic Excellence of Universities (89-E-FA04-1-4) and National Science Council (project 92-2213-E-007-018), Taiwan, R.O.C.

References

[1] K.K. Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 20, no. 1, pp. 39-51, 1998.

[2] P. Viola and M. Jones, "Rapid Object Detection Uaing Boosted Cascade of Simple Features," *Proc. Conf. Computer Vision and Patter Recognition*, pp. 511-518, 2001.

[3] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 20, no. 1, pp. 23-38, 1998.

[4] H.A. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection," *Proc. Conf. Computer Vision and Patter Recognition*, pp. 38-44, 1998.

[5] B. Heisele, T. Poggio, and M. Pontil, "Face Detection in Still Gray images," A.I. memo AIM-1687, Artificial Intelligence Laboratory, MIT, 2000.

[6] M.H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 24, no. 1, pp. 34-58, 2002.

[7] A. Pentland, B. Moghaddam, and T. Stanrner, "View-Based and Modular Eigenspaces for Face Recognition," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 84-91, 1994.

[8] A.L. Yuille, "Deformable Templates from Face Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 59-70, 1991.

[9] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machine: An Application to Face Detection," *Proc. Conf. Computer Vision and Patter Recognition*, pp. 130-136, 1997.

[10] B. Heisele, T. Serre, M. Pontil, and T. Poggio, "Component-based face detection, "*Proc. Conf. Computer Vision and Pattern Recognition*, pp. 657-662, 2001.

[11] H. Schneiderman and T. Kanade, "A statistical Method for 3D Object Detection Applied to Faces and Cars," *Proc. Conf. Computer Vision and Patter Recognition*, pp. 746-751, 2000.

[12] C. Liu "A Bayesian Discriminating Features Method for Face Detection" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 25, no. 6, pp. 725-740, 2003.

[13] A.J. Colmenarez and T.S. Huang, "Face Detection With Information-Based Maximum Discrimination," *Proc. Conf. Computer Vision and Patter Recognition*, pp. 782-787, 1997.

[14] Yoav Freund and Robert E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Computational Learning Theory: Eurocolt '95*, pp 23 -37, 1995

[15] M.-H. Yang, D. Roth, N. Ahuja, "A SNoW-based face detector." *Adbances in Neural Information Processing System 12*, pp 855 -861, MIT press, 2000.

[16] M. Alvira and R. Rifkin "An Empirical Comparison of SNow and SVMs for Face Detection" *CBCL Paper #193 / AI Memo #2001-004* ,MIT 2001

[17] R.-L. Hsu, M. Abdel-Mottaleb, A. K. Jain "Face Detection in Color Images" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 24, no. 5, pp. 696-706, 2002.

[18] C. Garcia and G. Tzorotas "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis" *IEEE Trans. Multimedia*, vol 1, no. 3, pp. 264-277, 1999.

[19] S.H. Huang and S.H. Lai "Readl-Time Face Detection in Color Video" Proceedings".10th International Multimedia Modelling Conference, , Pp:338 – 345, 2004